Knowledge-Based Systems 64 (2014) 44–58

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Experiment on sentiment embedded comparison interface

Li Chen *, Feng Wang, Luole Qi, Fengfeng Liang

Department of Computer Science, Hong Kong Baptist University, Hong Kong, China

ARTICLE INFO

Article history:
Received 8 August 2013
Received in revised form 13 February 2014
Accepted 23 March 2014
Available online 3 April 2014

Keywords:
Consumer reviews
E-commerce
Product comparison
Sentiment analysis
Feature-based review summarization
User study

ABSTRACT

Because the large amount of product reviews has been appearing in the current e-commerce sites, it becomes increasingly important to summarize these reviews, so as to support online buyers' information-seeking and decision-making process. However, little work has investigated how to present the sentiment information (as extracted from reviews) on the user interface, especially in the interface of supporting users to compare products. In this manuscript, we design three alternative sentiment-embedded comparison interfaces based on popular techniques, respectively called *opinion table*, *opinion bar chart*, *opinion cloud*. We then report results from two user studies on the developed interfaces. The first user study verified (1) the important role of comparison matrix in users' decision process, (2) the benefit of incorporating reviews into the comparison interface, and (3) the positive effect of showing features' sentiment info on aiding users' product comparison. Motivated by the first study's results, we performed the second user study to in depth compare the three alternative designs empirically. It turns out that the opinion bar chart, that mainly visualizes numerical feature sentiment scores via bars and qualitative adjective words via tool tip window, achieved significantly higher user assessments in terms of perceived information sufficiency, perceived ease of use and perceived cognitive effort. Users also behaved more active in opinion bar chart by manipulating the extracted features while less frequently viewing the raw textual reviews. The opinion cloud, that primarily visualizes the feature-associated opinion words in form of adjusted tags, was shown with better performance than opinion table, but slightly lower favor than opinion bar chart. In addition, this study revealed the effectiveness of showing *opinion features* (i.e., features with sentiment) in allowing users to examine the similarity and contrast across multiple products, and hence enabling them to make an informed and confident decision at the end.

1. Introduction

Product reviews that popularly appear in e-commerce sites (such as Amazon.com, Epinions.com) have been recognized as valuable information to assist new buyers in making purchase decision [13,26,55]. However, it is not easy for a new buyer to quickly grasp all reviews because: (1) the reviews were freely written in texts so they likely contain verbose and meaningless words; (2) the appearance of large amount of reviews often causes information overload to readers. Therefore, in recent years, there are increasing attentions paid to automatically extract opinions from reviews, which technique is called sentiment analysis (or opinion mining) [42,33,54], but most of existing works have emphasized algorithm development while less on user interface design and evaluation. Indeed, effectively presenting the extracted sentiment information is equivalently important, especially when users are performing product comparison.

According to researches in the area of marketing and decision psychology [44,4,45,21], *product comparison* is a crucial decision stage that buyers usually perform right before they make a choice. To be specific, after buyers screen down the number of available alternatives to a reduced consideration set, they often need to in-depth compare the selected candidates so as to decide on the final choice. To facilitate users to easily make the product comparison, the current e-commerce systems popularly provide a so called *Comparison Matrix* (CM) (see Fig. 1). It is basically in the form of an products (column) × features (row) matrix, by which buyers can make side-by-side comparison of products in terms of their attribute values. Previous studies claimed that such matrix can allow users to compare products more efficiently and accurately, because people do not need to retain the specific features' information about relevant candidates in their memory [2]. As a consequence, buyers can make better decision with less effort consumption [21]. However, the standard approach to create the

* Corresponding author. Address: 224 Waterloo Road, Kowloon Tong, Kowloon, Hong Kong, China. Tel.: +852 34117090; fax: +852 34117892.
E-mail addresses: lichen@comp.hkbu.edu.hk (L. Chen), fwang@comp.hkbu.edu.hk (F. Wang), llqi@comp.hkbu.edu.hk (L. Qi), ffliang@comp.hkbu.edu.hk (F. Liang).

http://dx.doi.org/10.1016/j.knosys.2014.03.020
0950-7051/

SPECIFICATION	Canon 7D	Nikon D300s	Pentax K-7	Sony a550
				
Megapixels	18.0 mp	12.3 mp	14.6 mp	14.2 mp
Sensor Size	22.3 x 14.9 mm	23.6 x 15.8mm	23.4 x 15.6mm	23.5 x 15.6mm
Sensor Type	CMOS	CMOS	CMOS	CMOS
File Size	5,184 x 3,456	4,288 x 2,848	4672 x 3104	4592 x 3056
Fps	8	7	5.2	7
ISO Range	100-6400	100-6400	100-3200	200-12800
AF Points	19 Points	51 Points	11 Points	9 Points
HD Video	1080i @ 30fps	720p @ 24fps	1080i @ 30fps	None
Live View	Yes	Yes	Yes	Yes
LCD Size	3.0	3.0	3.0	3.0
LCD Pixel Count	920,000	921,000	921,000	921,600
View Finder Coverage	100%	100%	100%	90%
Weatherproof	Yes	Yes	Yes	No
AEB / Frames	3	2 to 9	3 or 5	3
Built-in Flash	Yes	Yes	Yes	Yes
Shots Per Battery	800	950	980	950
Media Type	CF	CF & SD	SD	SD
Weight with Battery	2lbs	1.8lbs	1.6lbs	1.4lbs
Camera Specific Features	Nothing Worth Mentioning	Dual Memory Slot In-camera Editing	HDR In-camera Button for RAW In-body Image Stabalization Capable of Operating at -10C DNG Raw Mode	Tiltable LCD HDR In-camera In-body Image Stabalization Dual Memory Slot

Fig. 1. Example of a standard comparison matrix (CM).

comparison matrix is limited, because it only shows the features' static specifications [34]. Little innovation has been seen in incorporating the reviews' sentiment analysis results into it, though some studies stated that people tend to seek for other consumers' opinions on features during the process of product comparison [11].

Thus, in this manuscript, we are interested in evaluating the sentiment-embedded comparison matrix. In the following, we will use **opinion features** to denote features extracted from textual reviews, and **static features** to denote fixed specifications. The inherent advantage of opinion feature is that it can be associated with various types of information: (1) qualitative information such as the subjective words or phrases reviewers used to express the opinions (e.g., "good" in "the picture quality is good", and "easy to use" in "buttons on the menu are easy to use") [59]; (2) quantitative information such as the *opinion score* and *occurrence frequency*. The opinion score concretely represents the average opinion on a feature by aggregating all reviews to a product. The occurrence frequency indicates the amount or percent of reviewers who hold the corresponding positive (or negative) opinion. The critical issue is then how to well present all kinds of information on the user interface.

In our work, we have first developed three alternative sentiment-embedded CM interfaces, respectively based on bar chart, tag cloud, and tabular format. In bar chart based CM (that we call *opinion bar chart*), the numerical opinion scores are visualized via bars while the other information like qualitative adjective words are shown in a pop-up tool tip window. In tag cloud based CM (called *opinion cloud*), the adjective-noun pairs are emphasized in form of adjusted tags, while the opinion scores are kept in companion. The third interface (called *opinion table*) does not lay particular emphasis on quantitative or qualitative data, since both types of data are shown in parallel in a tabular format. In addition, each interface is interactive in nature, since users can freely manipulate any features by adding or removing them for comparison.

We have then conducted two user studies. In the first study, through both objective and subjective measurements, we not only verified the important role of comparison matrix in supporting users' decision making process, but also demonstrated the benefit of incorporating reviews into the comparison matrix and furthermore the significantly higher merit of opinion features relative to static features. Motivated by the first experiment's results, we carried out the second user study that empirically compared the three sentiment-embedded CM interfaces, i.e., *opinion bar chart*, *opinion cloud* and *opinion table*. The results show that the opinion bar chart achieves higher assessment scores in respect of most of subjective metrics. Particulary, its differences from opinion table reached at significant level, especially regarding users' perceived information sufficiency, perceived ease of use, and perceived cognitive effort. Users also behaved more active in opinion bar chart, while less frequently viewing the raw textual reviews. The opinion cloud was also found with better performance than opinion table, but slightly